Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 100 – 108

Procedia Computer Science

International Conference on Computational Modeling and Security (CMS 2016)

A Technique to Preserve Edge Information in Single Image Super Resolution

Amisha J Shah[a*], Suryakant B Gupta[b]

[a]*Assistant Professor,C.K.P.C.E.T.,Surat, India.*
[b]*FCIPT, Institute for Plasma Research, Gandhinagar, India*

Abstract

Goal of image super resolution is to enhance the size of an image without upsetting the inherited information. The quality of an enhanced image is conserved if the information around all kind of edges is preserved. The proposed novel approach potted the information around curvature edges and hard edges (abrupt transition in intensity) using Non Sub-Sampled Contourlet Transform (NSCT) based learning process. Furthermore the smoothness of smooth edges (gradual transition in intensity) is preserved by using soft edge smoothness prior as a regularizing parameter. The validity of the proposed approach is proven through simulation on several images.

Keywords: Image Super Resolution, Non Sub Sampled Contourlet Transform, Soft edge smoothness prior

1. Introduction

Improvement of visual information for human interpretation is the principle requirement in satellite imaging, medical science and surveillance etc. The cameras used at the on-board circuit of satellite are of low resolution because of limitation of weight and cost. Image super resolution (SR) becomes an important task to interpret the information carried by an image captured by these low resolution cameras. More over high quality images have also a concern in High Definition Television (HDTV). Thus image super resolution is a process of achieving the best image quality through single low resolution (LR) image or multiple low resolution images of the same scene. The super resolution approach offers benefit of utilization of the existing available low resolution imaging system.

* Corresponding author. Tel.:+91-9377616157
*E-mail address:*amisha.shah@ckpcet.ac.in

doi:10.1016/j.procs.2016.05.186

Traditionally, image super resolution is classified as multi image super resolution and single image super resolution. Multi image SR technique uses multiple images of the same scene. The fusion of information from multiple LR images enables the reconstruction of high resolution (HR) image easier as compared to single LR image. This is because multiple LR images contain different information and this additional information can be exploited to obtain a HR image. Thus it remains a challenging task to reconstruct a HR image from single LR image with less amount of information at hand. This paper proposes a novel learning based approach for single image super resolution.

Image super resolution techniques can be mainly categorized as reconstruction based techniques and learning based techniques. The reconstruction based SR approach estimate a HR image from several LR images using a regularizing parameter called prior. If a prior probability distribution on the super resolution image is available then this information may be used to regularize the estimation. Numerous generic smoothness priors are proposed by previous researchers[1, 2, 3, 4]. These generic smoothness priors offer gradient field with lower magnitude which limits the magnification factor. An adaptive regularization method is fused with an adaptive sparse domain selection to obtain a HR image[5]. In learning based approach, the relationship between an LR image and its corresponding HR image is examined via a pair of LR and HR patches. The training data is used to predict the higher resolution image. Xie Qinlan et al. proposed an example based single image super resolution method that classifies the high frequency patches of low resolution image into different classes[6]. For the edge areas of the LR image, the routine example based image super resolution algorithm can be used to implement the local and fine super-resolution. For the flat regions of the LR image, only interpolation algorithm is used for super resolution. Local high frequency details are synthesize using both ordinary and residual training data sets[7]. To reconstruct high frequency details in estimated HR image, Jian Zhang et al.[8] had also used ordinary and residual high frequency dictionary learning via sparse representation method. Pulak Purkait et al.[9] had used sparse dictionary learning technique in to adaptively select a prior. Here, HR image is estimated from a LR image by selecting a prior locally using knowledge of local statistics of LR image. Coupled dictionary training method, proposed by Jianchao Yang et al. is optimized as bilevel optimization problem[10]. As a bilevel optimization the author achieved optimization that includes l_1-norm minimization problem.

The algorithm proposed by Dai S.[4] integrates bilateral filtering into back projection method. The basic idea is that pixels which are nearby both in space domain and feature domain are passed through smoothing process. This may tend to blur the edges and the details around the edges may be lost. The concept of nearest neighbor based algorithm is also used to reconstruct a HR image[11, 12]. The reconstruction results of these methods are blur since the number of neighbor ‘k’ is fixed which tends to scrub out the edge details. The proposed work prioritizes the information around the edges while performing super resolution.

The paper, presented here, proposed a novel approach in which NSCT based learning algorithm is framed together with edge smoothness prior to preserve edge information during super resolution process. The training dataset is constructed by performing three level pyramidal decomposition of each image in the training set. The benefit of directionality offered by NSCT is obtained through the two levels directional decomposition at each pyramidal level. However, learning is performed between two coarser NSCT levels of a low-resolution image and training dataset. During the learning process a LR patch may map to multiple HR patches which leads to appearance of unwanted outliers in the super resolved image. Therefore, in the proposed approach a robust Lorentzian error norm is used to remove these outliers in contrast to minimum absolute difference criterion (MAD)[13]. The directional Gabor filter bank increases quality of multi directional images at the cost of complexity[13]. Furthermore, Gabor filter is a directional filter and hence it is not able to retain smoothness of smooth edges. Hence the proposed method uses a edge smoothness prior to maintain smooth edges of the test image in the super resolution process. Finally an objective function formed by considering global data term (residual between original image and estimated HR image) and a prior term. This objective function is optimized via Iterative Back Projection (IBP) method. The proposed approach yields better results considering both smoother regions as well as texture regions. The obtained results are compared via both qualitatively and quantitatively with the state-of-the-art results. The quantitative comparisons are done on the same platform.

The rest of the paper is organized as follow: The basic of Non Sub Sampled Contourlet Transform (NSCT) is described in section 2. Section 3 depicts robust estimation of NSCT coefficients at the finer scale of unknown HR image whereas section 4 explains the global constraint and removal of edge artifacts using soft edge smoothness prior. Through simulation on variety of images, the validation of the proposed algorithm is reported in section 5.

2. Fundamentals of Non Sub Sampled Contourlet Transform

The visual information at discontinuities, mainly at edge points and along the contours may not be smooth in the reconstructed HR image. Wavelet and Curvelet are the well known transforms that offers good smoothness at discontinuities. But the wavelet is limited to capture the information along the horizontal direction, vertical direction and the diagonal direction (the direction at an orientation of 45^0). It is also not able to offer smoothness along the contours. Other side, Curvelet overcomes this limitation. It is implemented using a rotation operation and a 2-D frequency partition based on the polar coordinates. Therefore it is simple to implement in frequency domain but difficult in discrete domain[14].The Contourlet transform offers high directionality but it is shift variant. However, image analysis applications such as edge detection, image enhancement requires a transform that is shift invariant. NSCT is a shift invariant transform that offers high directionality[15]. The NSCT constitute of multiscale (pyramidal) as well as multidirectional decomposition of an image. The multiscale decomposition property of NSCT decomposes an image into different frequency bands namely low frequency band and high frequency band. The image, containing low frequency band is used as an input image for further pyramidal decomposition whereas non sub sampled directional filter bank (NSDFB) is applied to high frequency band image to obtain multidirectional decomposition. Directional decomposition by NSCT will generate the sub bands in power of 2. If "*n*" level directional decomposition is done, then NSCT will generate 2^n sub bands. For more understanding, NSCT decomposition of an image by considering one level pyramidal decomposition and one level directional decomposition at the same level is shown in Fig.1(a). Directional sub bands for three level directional decomposition are shown in Fig.1(b).

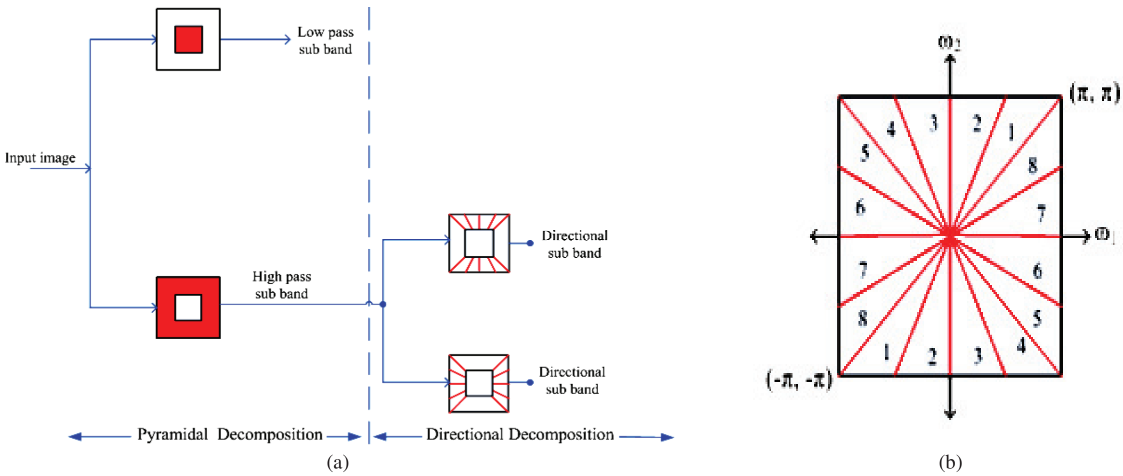


Fig. 1. (a)NSCT decomposition for one level pyramidal decomposition along with one level directional decomposition, (b) Directional sub bands of a high pass sub band image for three level directional decomposition within $+\pi$ to $-\pi$ range

3. Robust estimation of NSCT coefficients

Single image super-resolution via learning in NSCT domain is based on the concept that, in a non sub-sample pyramid, every coefficient at the coarser level can be related to the coefficient at the next finer level for the similar orientation[17]. To estimate the NSCT coefficients of high-resolution image via learning process, the training dataset is to be constructed first. Here the training set is constructed by considering the NSCT coefficients of all the high resolution images of the database. These NSCT coefficients of the training set are obtained by decomposing each image of database into three pyramidal level and four directional sub-bands at each pyramidal level. The given low resolution test image is decomposed into two level pyramidal decomposition along with two level directional decomposition at each pyramidal level. The schematic diagram of NSCT decomposition of test image and one of the training images is shown in Fig.2.

The four directional sub-bands related to the third pyramidal level of the test image are to be learned in NSCT domain. The learning is done by searching the best match for NSCT coefficients corresponding to the two coarser pyramidal levels of the given test image from the training database (considering NSCT coefficients related to only two coarser pyramidal levels of each training image). The performance of learning based super resolution strongly depends on HR coefficients retrieved from the training data for input LR coefficients. Therefore, to obtain a HR image, it is important to properly estimate HR coefficients. Hence, a robust statistic is adopted to estimate NSCT coefficients of four directional sub-bands related to the third pyramidal level of the up sampled test image.

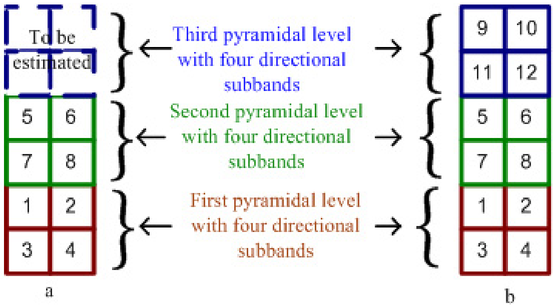


Fig. 2. NSCT decomposition of (a) test image (b) one of the training images.

Learning in NSCT domain starts with partition of two coarser directional levels of the up sampled test image into 4×4 block of NSCT coefficients. As each pyramidal level is decomposed into four directional sub-bands, eight directional sub bands for two coarser pyramidal levels are obtained. Denote 4×4 block of NSCT coefficients from each directional sub band as B_l, where l = 1, 2, ----, 8. In support of this set of 128 ((4x4) block*8 sub bands) NSCT coefficients, perform a search for the best match from two coarser pyramidal levels of each training image as follow:

$$K = \underset{k}{argmin}[d_1 + d_2 + ... + d_8] \tag{1}$$

Where, $k = 1, 2,----, N$ training images. K is the k^{th} training image presenting best match for the test image and d_l is the error defined as

$$d_l = \rho(B_l - T_l^k, \sigma) \tag{2}$$

Where, l = 1, 2, ----, 8, $\rho(\cdot)$ is a robust error norm, σ is a scale parameter and T_l^k denotes 4x4 block of l^{th} sub-band of k^{th} image from training set . Robust estimation of NSCT coefficients highly depends on selection of error norm $\rho(\cdot)$. To analyse the behaviour of the error norm, consider its derivative function $\psi(\cdot)$. If most commonly used least square error, defined as e^2/σ^2, is selected as $\rho(\cdot)$ function for error e then the selected best match may be an outlier. This is because the derivative function of least square error norm[18] increases linearly with error e as shown in Fig.3(b). Hence, in contras to MAD[13] the Lorentzian error norm is adopted as a robust error norm. The derivative function of the Lorentzian error norm implies that if e increases beyond a fixed value determined by the scale parameter σ, the influence of outlier is reduced as shown in Fig.3(d).

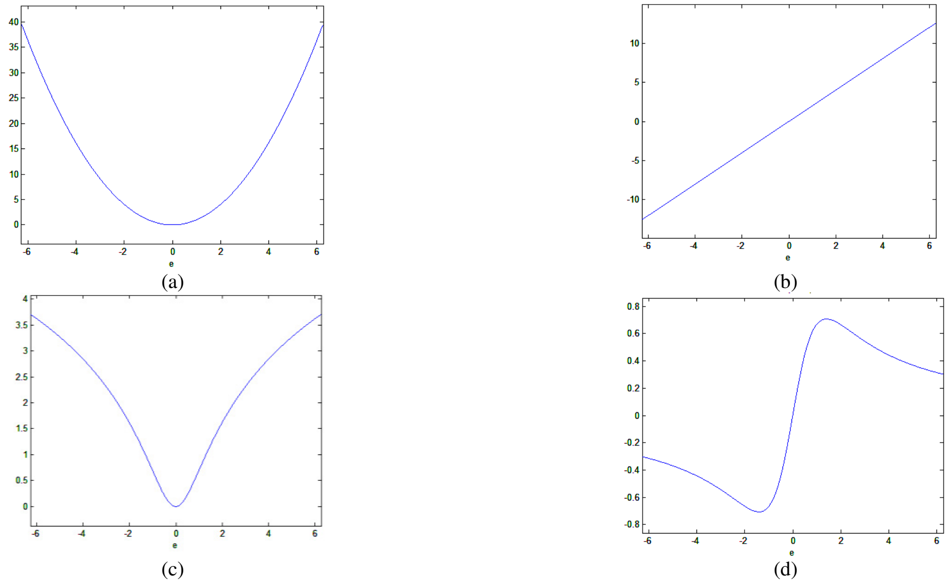


Fig. 3. (a) ρ (·) function and (b) ψ (·) function for least square error norm whereas (c) and (d) represents ρ (·) function and ψ (·) function for Lorentzian error norm respectively.

The Lorentzian error norm is described as:

$$\rho(e,\sigma)=log\left[1+\frac{1}{2}\left(\frac{e}{\sigma}\right)^2\right] \tag{3}$$

Thus d_l in equ.(1) is calculated by considering the Lorentzian error norm defined in equ.(3). From the best matched training image obtained through equ.(1), get the unknown NSCT coefficients (in form of 4×4 block) in the sub-bands 9-12 as follow:

$$B_9 = T_9{}^K, B_{10} = T_{10}{}^K, B_{11} = T_{11}{}^K, B_{12} = T_{12}{}^K \tag{4}$$

This has to be repeated for each 4×4 block corresponding to eight directional sub-bands of the test image. At last, the test image with three levels pyramidal decomposition is obtained. The inverse transform of this image results in a HR image that retrives informations around curved edges as well as hard edges. Smooth edges of this resultant HR image are retrived by superimposing smooth edge prior[19].

4. Global constraint and removal of edge artifacts

The high resolution image obtained through learning in NSCT domain increases visual information along abrupt discontinuities and contours present in the test image. Visual information at smooth edges can be improved by using a smooth edge prior as a regularizing parameter. The smooth edge prior $|I|g$ can be defined as[20]:

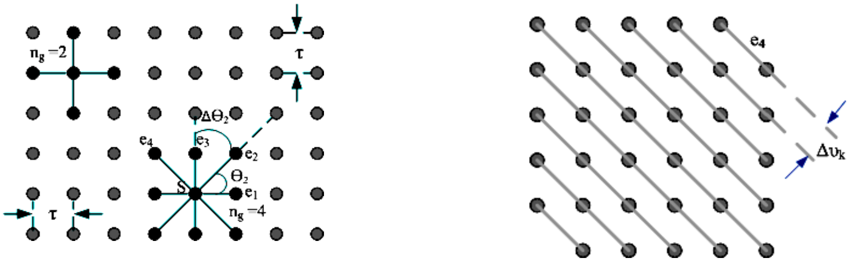
$$|I|_g = \sum_{k=1}^{n_g}\left(w_{pq}\sum_{(p,q)\in N_k}\left|I_p - I_q\right|\right) \tag{5}$$

Where, w_{pq} is the weight of an edge connecting neighbourhood pixels p and q of an image I. N_k represents family of edge lines for each edge vector e_k of an edge set. The edge set can be represented as a set of vectors $\{e_k|1\leq k\leq n_g\}$ where n_g defines the order of the neighbourhood system [21]. Each vector in a set is ordered by its corresponding orientation angle θ_k such that $0\leq \theta_1\leq \theta_2\leq \theta_3\leq....\leq \theta n_g \leq \pi$. As an example, set of vectors for $n_g = 2$ and $n_g = 4$ are shown in Fig.4(a). Each vector e_k has a family of edge lines N_k having same interline angle θ_k and interline distance Δv_k as shown in Fig.4(b).

The edge weights for k^{th} edge vector family can be defined as[21]:

$$w_k = \frac{\tau^2 \Delta\theta_k}{2\left|e_k\right|} \tag{6}$$

Where, $|e_k|$ is the Euclidean length vector, τ is the interval of the grid on the weighted grid graph and $\Delta\theta_k=\theta_{k+1} - \theta_k$.



(a) (b)

Fig. 4. (a) Neighbourhood system for $n_g = 2$ and $n_g = 4$, (b) Family of edge lines for vector e_4.

On adding equ.(5) as a regularization term to a global constrain term, a cost function is formed. The cost function thus can be represented as:

$$E = \left[I_l - d(I_H) \right] + \lambda \left[\left| I_l \right|_g - \left| d(I_H) \right|_g \right] \tag{7}$$

Where, I_l is the LR image. I_H represents sharp initialization obtained through the learning process. $[I_l - d(I_H)]$ is the global constrain term and $[|I_l|_g - |d(I_H)|_g]$ is a regularization term (prior term). $|I_l|_g$ and $|d(I_H)|_g$ are obtained using equ.(5). Parameter λ is the step size that controls the trade off between the global constraint term and prior operator. Final HR image is obtained by minimizing 'E' through an Iterative Back Projection algorithm. The order of neighbourhood system defined in smooth edge prior affects the quality of reconstructed HR image. The edges are smoother with higher value of n_g. This is demonstrated in Fig.5.

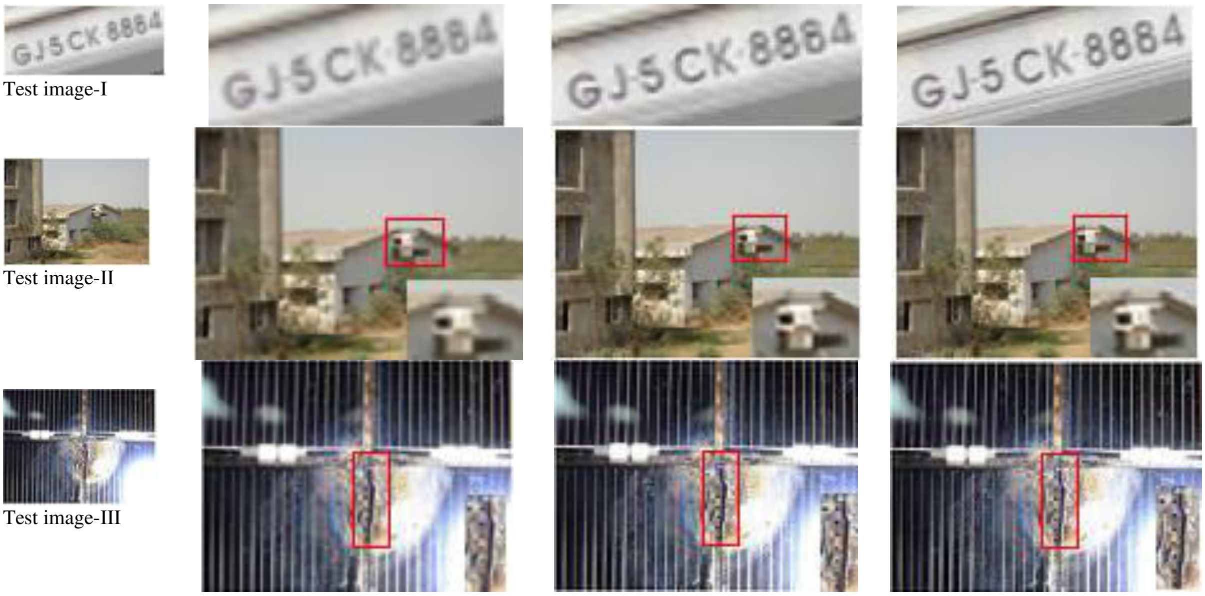
LR image1

(a) (b) (c)

Fig. 5. (a) Test image, (b) and (c) depicts SR result (x2) with $n_g = 2$ and $n_g = 8$ respectively.

5. Results and Discussion

The proposed algorithm starts with the construction of the training database which needs to calculate the NSCT coefficients of each training image for three levels pyramidal decomposition along with four directional sub-bands at each level. Throughout the experiment same training data set is used. The training data set consists of 100 images of different classes like natural images, texture images and facial images etc. The proposed novel approach uses learning in NSCT domain which preserved the information along the hard edges and curvature edges. A smoothness prior with $n_g = 8$ is applied to the image obtained through the learning process. The parameter λ is set to 0.001 as edges are sharpened due to learning in NSCT domain. The optimization is obtained through IBP method with 12 iterations only.



(a) (b) (c) (d)

Fig. 6. (a) LR image, SR result (x2) using (b) Bicubic interpolation (c) J. Yang et. al[22] (d) proposed approach.

The proposed method is verified on different types of images like natural images, satellite ESD images, text images, texture images etc. As a text image the proposed algorithm is applied to a license plate (test image I) that is captured through a camera used for vehicle surveillance. The proposed approach also gives satisfactory result for an image of a satellite solar coupon exposed to a permanent sustained arc as shown in Fig.6(test image III).

The results obtained by the proposed approach are compared with Bicubic interpolation, J. Yangs approach[22, 23] and K.Kim[24] by means of perception as well as qualitatively. The proposed algorithm shows a substantial improvement. Perception comparison, demonstrated in Fig.6 through Fig.8, shows minimum edge artifacts for the proposed approach. The quantitative measurement is done in table 1 using peak signal to noise ratio (PSNR) and feature similarity (FSIM) index[25]. With a glance on comparative table 1, the proposed approach offers desired results like improved PSNR and higher FSIM value. PSNR improvement of more than 1.5 dB (with respect to Bicubic interpolation) is observed for the test image II and test image III as these images contain edges of different orientations. For test image I, PSNR is improved by 10.6 dB with respect to Bicubic interpolation because this image contains curve edges.

Table1. PSNR and SSIM COMPARISON FOR ZOOMING FACTOR 2

Source Image	Bicubic Interpolation		SR via J. Yang et. al[22]		Proposed Approach	
	PSNR	FSIM	PSNR	FSIM	PSNR	FSIM
Test Image I	25.8325	0.8835	27.2115	0.9108	36.4582	0.9875
Test Image II	31.2447	0.9233	32.9040	0.9500	33.6731	0.9565
Test Image III	21.7285	0.8148	22.5101	0.8539	22.9861	0.8477

The result for zooming factor 3 is compared with another learning based algorithm[23] in Fig.7. The proposed algorithm produces more natural result than [23].

PSNR=34.3872 PSNR=35.5611

(a) (b) (c)

Fig. 7. (a) LR image, SR results (x3) using (b) J Yang et. al[23], (c) proposed algorithm

Finger print is one of biometric trait. An image of finger print with high resolution plays a vital role in forensic investigation. A finger print image is zoomed by a factor of 4 as shown in Fig.8.

PSNR= 18.2501 PSNR= 20.2913 PSNR= 21.0401

(a) (b) (c) (d)

Fig. 8. (a) LR image of a fingerprint image, SR results for zooming factor 4 (b) Bicubic, (c) K. Kim et. al[24], (d) Proposed approach.

Conclusion

The proposed method is useful when one has to use single observed image to improve its resolution. The results obtained for different types of images prove validity of the proposed approach. The proposed learning based algorithm is based on the fact that every coefficient at the coarser level of the Non Sub-Sample Contourlet Transform can be related to the coefficient at the next finer level for the similar orientation. The probable outliers in learning process are rejected by adopting robust Lorentzian error norm. The image obtained through this learning process preserves information along the hard edges as well as along the curve edges presented in the image. The smoothness of smooth edges is conserved using a smooth edge prior. The reported results demonstrate better resolution with minimum complexity.

References

1. Sina Farsiu, M. Dirk Robinson, Michael Elad, Peyman Milanfar. Fast and Robust Multiframe Super Resolution. IEEE transactions on image processing 2004; vol.13, no.10, p.1327-1344.
2. David Capel , Andrew Zisserman. Super-resolution Enhancement of Text Image Sequences. IEEE Comput.Soc.ICPR 2000; vol. 1, no. 1, p. 600-605.
3. Xueting Liu, Daojin Song, Chuandai Dong, Hongkui Li. MAP-Based Image Super-resolution Reconstruction. World Academy of Science, Engineering and Technology 2008; vol.37, p. 208-211.
4. Shengyang Dai, Mei Han, Ying Wu, Yihong Gong. bilateral back-projection for single image super resolution. IEEE Proceedings of ICME 2007; p. 1039-1042.
5. Weisheng Dong, Lei Zhang, Guangming Shi, Xiaolin Wu. Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization. IEEE transactions on image processing 2011; vol. 20, no. 7, p. 1838-1857.
6. Xie Qinlan, Chen Hong, Cao Huimin. Improved Example-based Single-image Superresolution. IEEE CISP 2010; p.1204-1207.

7. Changhyun Kim, Kyuha Choi, Jong Beom Ra. Improvement On Learning-Based Super-Resolution By Adopting Residual Information And Patch Reliability. IEEE ICIP 2009; p.1197-1200.
8. Jian Zhanga, Chen Zhaob, Ruiqin Xiongb, Siwei Mab, Debin Zhaoa. Image Super-Resolution Via Dual- Dictionary Learning And Sparse Representation. IEEE International Symposium on Circuits and Systems (ISCAS) 2012; p. 1688-1691.
9. Pulak Purkait, Bhabatosh Chanda. Image upscaling using multiple dictionaries of natural image patches. Asian Conference on Computer Vision, 2012.
10. Jianchao Yang, Zhaowen Wang, Zhe Lin, Scott Cohen, Thomas Huang. Couple Dictionary Training for Image Super-resolution. IEEE transactions on image processing 2012; vol.21, no. 8, p. 3467-3478.
11 Chan T M, Zhang J, Pu J, Huang Hua. Neighbor embedding based super-resolution algorithm through edge detection and feature selection. Pattern Recognition Letters 2009; 30(5), p. 494–502.
12 Bevilacqua M, Roumy A, Guillemot Ch, Marie-Line Alberi Morel. Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding. British Machine Vision Conference 2012; p. 1-10.
13 Amisha J. Shah, Rujul Makwana, Suryakant B. Gupta. Single Image Super-Resolution via Non Sub-sample Contourlet Transform based Learning and a Gabor Prior. International Journal of Computer Applications 2013; vol.64, no.18, p. 32-38.
14 Minh N. Do, Martin Vetterli. The Contourlet Transform: An Efficient Directional Multi resolution Image Representation. IEEE transactions on image processing 2005; vol.14, no.12, p. 2091-2106.
15 Yue Lu, Minh N. Do. A New Contourlet Transform With Sharp Frequency Localization. IEEE Proceedings of International Conference on Image Processing 2006; p. 1629-1632.
16 Jianping Zhou, Arthur L. Cunha, Minh N. Do. Non subsampled Contourlet Transform: Construction And Application In Enhancement. IEEE transactions on image processing 2006; vol. 15, no. 10, p. 469-472.
17 C. V. Jiji, Subhasis Chaudhuri. Single-Frame Image Super-resolution through Contourlet Learning. EURASIP Journal on Applied Signal Processing 2006; vol.2006, Article ID 73767, p. 1-11.
18 Michael J. Black, Guillermo Sapiro, David H. Marimont, David Heeger. Robust Anisotropic Diffuusion. IEEE transactions on image processing 1998; vol. 15, no. 3, p. 421-432.
19 Yuri Boycov, Vladimir Kolmogorov. Computing Geodesics and Minimal Surfaces via Graph Cuts. Proceedings of the Ninth IEEE International Conference on Computer Vision 2003; vol. 2.
20 Dai S, Han M, Xu W, Wu Y, Gong Y. SoftCuts: A Soft Edge Smoothness Prior for Color Image Super-Resolution. IEEE transactions on image processing 2009; vol. 1, no. 5, p. 969-981.
21 Boykov Y, Veksler O, Zabih R. Fast Approximate Energy Minimization via Graph Cut. IEEE transactions on pattern analysis and machine intelligence 2001; vol.23, no. 11, p. 1222-1239.
22 J. Yang, John Wright, Thomas Huang, Yi Ma. Image super-resolution via sparse representation. IEEE Transactions on Image Processing 2010; vol.19, Issue 11, p. 2861-2873.
23 Jianchao Yang , John Wright, Thomas Huang, Yi Ma. Image Super-Resolution via Sparse Representation. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 2008; p. 1-8.
24 Kim K I, Kwon Y. Single-image super-resolution using sparse regression and natural image prior. IEEE Trans. Pattern Analysis and Machine Intelligence 2010; vol. 32, no. 6, p. 1127-1133.
25 Lin Zhanga, Lei Zhanga, Xuanqin Moub, David Zhanga. FSIM: A Feature Similarity Index for Image Quality Assessment. IEEE Transactions on Image Processing 2011; vol. 20, no. 8, p. 2378-2386.